A dynamic network in a dynamic population: asymptotic properties

Tom Britton,* Mathias Lindholm† and Tatyana Turova‡

January 18, 2013

Abstract

We derive asymptotic properties for a stochastic dynamic network model in a stochastic dynamic population. In the model, nodes give birth to new nodes until they die, each node being equipped with a social index given at birth. During the life of a node it creates edges to other nodes, nodes with high social index at higher rate, and edges disappear randomly in time. For this model we derive criterion for when a giant connected component exists after the process has evolved for a long period of time, assuming the node population grows to infinity. We also obtain an explicit expression for the degree correlation ρ (of neighbouring nodes) which shows that ρ is always positive irrespective of parameter values in one of the two treated submodels, and may be either positive or negative in the other model, depending on the parameters.

Keywords: Degree correlation, dynamic networks, phase transition, random graphs, stationary distribution.

2010 Mathematics classification subjects. Primary: 92D30, Secondary: 60J80.

1   Introduction

The models of dynamical graphs are defined by the rules of attachment and deletion of vertices and edges chosen to fit particular processes in nature. Intensive study in this area began in the end of the 20th century when a number of models were formulated primarily in the physics literature (see, e.g., Barabási *et al.* (1999), Callaway *et al.* (2001), Malyshev (1998)). The important paper by Bollobás, Janson and Riordan (2007) provided a unified approach based on branching processes to many known models of random networks (see the reference list in Bollobás *et al.* (2007))

In the present paper we continue the study of a Markovian model describing a random time-dynamic network in a random time-dynamic population. The model was originally defined by Britton and Lindholm (2010) extending an earlier model of Turova (2003),

---

*Department of Mathematics, Stockholm University, SE-106 91 Stockholm, Sweden, tomb@math.su.se

†Uppsala University, Uppsala, Sweden. Current address: AFA Insurance, Stockholm, Sweden, Mathias.Lindholm@afaforsakring.se

‡Mathematical Center, Lund University University, Sweden, tatyana@maths.lth.se

which in turn was derived from a general model of Malyshev (1998). The population process is a Markovian linear birth-and-death process in which individuals are assigned random i.i.d. social indices. Given the population process, a Markovian network is defined where edges between individuals appear and disappear in such a way that "social" individuals tend to have more neighbours. Notice also, that it was shown by Turova (2002) that the model without social index and also without deletion of nodes includes as a subcase yet another model studied by Callaway *et al.* (2001).

We study asymptotic properties of the network: what properties will the network (and population) have after having evolved for a long time assuming the size of the population has grown large. One can check that a snap-shot of the limiting network of Britton and Lindholm (2010) falls into the general class of inhomogeneous random graphs introduced by Bollobás, Janson and Riordan (2007). In particular, one can consult Britton and Lindholm (2010) as well as Bollobás, Janson and Riordan (2007) for the age, the type and the degree distribution in this model.

In the present paper we shall make essential use of the theory of Bollobás, Janson and Riordan (2007) to study the phase transition in the model by Britton and Lindholm (2010). More precisely, we will determine the critical values which separate the set of parameters under which the model with a high probability has a giant connected component (i.e. of order of the entire network), and the area of parameters which do not produce a giant component. This extends earlier results of Turova (2007) for the model without social index.

Besides global properties of the network, e.g. phase transition, we shall study here a characteristic, which tells more about the local structure, namely the degree correlation, or the mixing coefficient as it was introduced by Callaway *et al.* (2001). Inspiring numerical results on the mixing coefficient in different empirical networks were presented and analyzed already by Newman (2002). Recently Bollobás, Janson and Riordan (2011) derived formulae for the mixing coefficient in terms of small subgraphs counts for a rather general graph model, which includes, in particular, the inhomogeneous graphs.

According to empirical results provided by Newman (2002) real life networks appear to be of two classes: assortative, when the mixing coefficient is positive, these are primarily the social networks, or disassortative, when the mixing coefficient is negative, which is a feature of technological or biological networks. Still, most most of the models tend to have positive degree correlation/mixing coefficient. However, Newman (2002) could construct an example of a network model where the mixing coefficient changes sign depending on the parameters of the model. Lately Bollobás, Janson and Riordan (2011) gave another recipe to construct a network with a negative mixing coefficient. Therefore it is a challenge to look for an example of a somewhat "naturally grown" random network which possesses the property of enabling both positive and negative mixing coefficient depending on the parameters. As it is pointed out by Newman (2002) the property of assortativity affects qualitatively the sharpness of the phase transition with respect to the size of the giant component: the phase transition is sharper for the dissasortative networks.

Here we provide an example of a dynamic (growing) social network with a degree correlation (mixing coefficient) that may be either negative or positive depending on model parameters. We derive an explicit formula for the degree correlation, showing the dependence of the sign of the mixing coefficient on the parameters. Notice, that the formula for the mixing coefficient for the model of Callaway *et al.* (2001) derived by Newman (2002), is a subcase of the formula we prove here.

Although our model is a subcase of the general model studied in Bollobás, Janson and Riordan (2011), and hence the general formula from this paper should be applicable for our model as well, we find a somewhat more direct way to get a formula for the degree correlation in our case. We use the invariant measures for the random walk associated with the graph. It is worth noting that our method is not restricted to the particular model we study here, but would work as well for a class of inhomogeneous random graphs. However this will be a subject of a separate study.

2 The Markovian random network in a Markovian dynamic population

Below we define the Markovian random network in a Markovian dynamic population, originally defined in Britton and Lindholm (2010). There are two versions of the model, referred to as the Uniform (U) version and the Preferential (P) version. In what follows a network denotes a finite set of nodes (the population) together with undirected edges connecting pairs of nodes. Nodes that are directly connected by an edge are called neighbours (in sociological applications nodes correspond to individuals, edges to some type of friendship, and neighbours are referred to as friends). The model is dynamic in the double sense that nodes are born and may die, and the same applies to edges.

2.1 The model and its two versions

We first define the node population dynamics. Let $Y(t)$ denote the number of nodes alive at time t, and assume that $Y(0) = 1$. While alive, each node gives birth to new nodes at the constant rate λ and each node lives for an exponentially distributed time having mean $1/\mu$ (denoted $\mathrm{Exp}(\mu)$), so each node dies at the rate μ. We assume that $\lambda > \mu$ implying that the expected number of births during a life-span is larger than 1. This means that the node population is modelled by a Markovian super critical branching process. Additional to this, each node i is at birth given a random "social index" S_i having distribution F_S on $\mathbb{R}^+$, independent and identically distributed for different nodes. We will throughout assume that S has finite mean $\mathsf{E}[S] := \mu_S < \infty$.

There are two versions of the model for births and deaths of edges, both being Markovian given the node population and their social indices. In both versions, nodes are isolated at birth, i.e. have no neighbours. During the life of a node i, having social index $S_i = s$, it creates new neighbour edges at rate αs, also the same for both models. The difference between the two versions lie in how neighbour nodes are selected. In the uniform (U) version the neighbour node is chosen uniformly at random among all living nodes. In the preferential (P) version of the model the neighbour node is instead chosen at random among all living nodes with probabilities proportional to their social index. Finally, in both versions each edge is removed, independently of everything else, at the rate β. If a node dies all edges connected to the node in question are removed. We emphasize that a node gets new neighbours in two ways: it "creates" new neighbour edges itself but may also be selected as neighbour of another nodes that has created a neighbour edge.

2.2 Comments on the model

The model has four parameters: the birth and death rates of nodes, λ and μ respectively, and the death rate β of edges and α which is related to the birth rate of new edges. Beside these four parameters there is the distribution F_S for the social indices $\{S_i\}$ (in what follows we let f_s denote the corresponding density function).

The P-version of the model is inspired by the preferential attachment model [1], still being different in that here the probability of receiving edges is determined at birth whereas it is determined by random events during life in the preferential attachment model. In both models there is also an age-factor in the sense that older nodes tend to have more edges.

Some submodels are worth mentioning. One is where there is no node-heterogeneity and nodes have the same social index $S \equiv s$ (for example set to 1 without loss of generality). Loosely speaking, allowing node-heterogeneity makes it possibly to achieve degree distributions having heavy tails: if F_S is heavy tailed there will be some nodes with very high social indices that hence have a large number of neighbours (cf. the next section). The case where $\mu = 0$ and $S \equiv 1$ has been studied by Turova (2002) and (2007) who derive more results for this submodel.

2.3 Known results of the model

The model allows for multiple edges and self-loops which usually does not make sense in applications. However, it was shown by Britton and Lindholm (2010) that the *proportion* of such edges is asymptotically negligible as long as $\mathsf{E}[S] < \infty$ for the U-version, and as long as $\mathsf{E}[S^2] < \infty$ for the P-version of the model. As a consequence the network will then have identical properties if loops and multiple edges are ignored or not allowed.

In the rest of the paper we will only consider the case where the node population tends to infinity, i.e. we condition on the event $B := \{\omega; Y(t) \to \infty\}$, which has positive probability since it was assumed that $\lambda > \mu$. The first two results follow from standard branching process theory:

Asymptotic population size: As $t \to \infty$

$$Y(t) \sim e^{t(\lambda - \mu)}.$$

Stable age distribution: As $t \to \infty$ the age A of a randomly selected living node will satisfy $A \sim \text{Exp}(\lambda)$ having mean $1/\lambda$, see e.g. Section 3.4 in Haccou et al. (2005).

Social index distribution: Since the social indices are defined to be i.i.d. having distribution F_S, and they only affect the occurrence of edges, the social index of a randomly selected living node will have the same social index distribution F_S.

Asymptotic degree distribution: As $t \to \infty$ the degree (i.e., the number of neighbours) D of a randomly selected living individual has a mixed Poisson distribution in both versions of the model (Britton and Lindholm, 2010). For the U-version it is

$$D^{(U)} \sim \text{MixPo}\left(\frac{\alpha(S + \mu_S)\left(1 - e^{-(\beta + \mu)A}\right)}{\beta + \mu}\right), \tag{1}$$

where $A \sim \text{Exp}(\lambda)$ and $S \sim F_S$ are independent. This simply means that, conditional on $A = a$ and $S = s$, the degree is Poisson distributed with mean parameter equal to

$\alpha(s + \mu_S)\left(1 - e^{-(\beta+\mu)a}\right) / (\beta + \mu)$. The mean and variance of this distribution are given by

$$\mathsf{E}[D^{(U)}] = \frac{2\alpha}{\lambda + \beta + \mu}\mu_S,$$

$$\mathsf{V}[D^{(U)}] = \frac{2\alpha}{\lambda + \beta + \mu}\mu_S$$
$$+ \frac{4\lambda\alpha^2}{(\lambda + \beta + \mu)^2(\lambda + 2(\beta + \mu))}\mu_S^2 + \frac{2\alpha^2}{(\lambda + \beta + \mu)(\lambda + 2(\beta + \mu))}\mathsf{V}[S].$$

For the P-version of the model the degree distribution satisfies

$$D^{(P)} \sim \mathrm{MixPo}\left(\frac{2\alpha S\left(1 - e^{-(\beta+\mu)A}\right)}{\beta + \mu}\right), \tag{2}$$

where $A \sim \mathrm{Exp}(\lambda)$ and $S \sim F_S$ are independent. The mean and variance are given by

$$\mathsf{E}[D^{(U)}] = \frac{2\alpha}{\lambda + \beta + \mu}\mu_S,$$

$$\mathsf{V}[D^{(U)}] = \frac{2\alpha}{\lambda + \beta + \mu}\mu_S$$
$$+ \frac{4\lambda\alpha^2}{(\lambda + \beta + \mu)^2(\lambda + 2(\beta + \mu))}\mu_S^2 + \frac{8\alpha^2}{(\lambda + \beta + \mu)(\lambda + 2(\beta + \mu))}\mathsf{V}[S],$$

i.e. the same mean degree but a larger variance. For both versions it is seen that the variance of the degree distribution increases with the variance of the social index distribution, so having a heavy-tailed social index distribution F_S implies that also the degree distribution will be heavy-tailed.

3 Results

In what follows we derive some further properties of the model, still assuming $t \to \infty$ and conditioning on that the node population grows beyond all limits. We do this for the U-version and only present the results for the P-version which is obtained in a similar way.

3.1 Type-distribution of neighbour nodes

Assume t to be large and consider a randomly selected individual alive at this time having age a and social index s. We derive the distribution of the number of (a', s')-neighbours this (a, s)-node has. The number of neighbours of any type (a', s') will be Poisson distributed, so we first compute the mean of the distribution and then divide by the total mean thus obtaining the type-distribution.

How many neighbours of type (a', s') does the (a, s)-individual have? Any such edge must have been created at some time $\tau \in (t - a \wedge a', t)$ ('$\wedge$' denotes minimum: both individuals must have been born). At such τ our (a, s)-individual creates such an edge at rate $\alpha s \lambda e^{-\lambda(\tau - (t-a'))} f_S(s')$, because it creates edges at rate αs and the second part denotes

5

the fraction of nodes having age $\tau - (t - a')$ (at τ) and type s'. Similarly, it receives edges at the rate $\alpha s' \lambda e^{-\lambda(\tau-(t-a'))} f_S(s')$ because there are $Y(\tau)\lambda e^{-\lambda(\tau-(t-a'))} f_S(s')$ having the "correct" age and type, each of them creates edges at rate $\alpha s'$ and with probability $1/Y(\tau)$ it reaches our (a, s)-individual. Any such edge created at τ remains at t if and only if both the (a', s')-individual and the edge survive until t (we ave already conditioned on that our (a, s)-individual lives at t). The probability for this equals $e^{-(\beta+\mu)(t-\tau)}$. The total expected number of edges with (a', s')-individuals hence equals

$$m^{(U)}(a', s'|a, s)da'ds' = da'ds' \int_{t-a\wedge a'}^{t} \alpha(s + s')\lambda e^{-\lambda(\tau-(t-a'))} f_S(s')e^{-(\beta+\mu)(t-\tau)}d\tau. \quad (3)$$

$$= \alpha(s + s')f_S(s')\lambda e^{-\lambda a'}\frac{e^{(\lambda-\beta-\mu)a\wedge a'} - 1}{\lambda - \beta - \mu}da'ds'. \quad (4)$$

It is easy to check from this that the expected total number of neighbours of any type of the (a, s)-node, the integral of this quantity over all a' and s', equals

$$m^{(U)}(\cdot, \cdot|a, s) = \frac{\alpha(s + \mu_S)}{\beta + \mu}\left(1 - e^{-(\beta+\mu)a}\right). \quad (5)$$

This also confirms that the unconditional degree distribution has the mixed Poisson distribution specified in Equation (1).

As a consequence, for the U-version of the model, the type distribution of the neighbours of an individual of type (a, s), which equals the expected number of (a', s')-types divided by the total expected number of types, becomes

$$f^{(U)}(a', s'|a, s) = \frac{m^{(U)}(a', s'|a, s)}{m^{(U)}(\cdot, \cdot|a, s)} = \frac{(s + s')f_S(s')}{s + \mu_S}\frac{(\beta + \mu)\lambda e^{-\lambda a'}\left(e^{(\lambda-\beta-\mu)a\wedge a'} - 1\right)}{(\lambda - \beta - \mu)(1 - e^{-(\beta+\mu)a})}.$$

We hence have conditional independence: $f^{(U)}(a', s'|a, s) = f^{(U)}(a'|a)f^{(U)}(s'|s)$, where

$$f^{(U)}(a'|a) = \frac{(\beta + \mu)\lambda e^{-\lambda a'}\left(e^{(\lambda-\beta-\mu)a\wedge a'} - 1\right)}{(\lambda - \beta - \mu)(1 - e^{-(\beta+\mu)a})}, \quad (6)$$

$$f^{(U)}(s'|s) = \frac{(s + s')f_S(s')}{s + \mu_S}. \quad (7)$$

Admittedly, the notation here, and in what follows, is a bit sloppy in that $f^{(U)}(a'|a)$ and $f^{(U)}(s'|s)$ really denote different density functions, as indicated by the difference in arguments.

Using similar arguments, but for the P-version of the model, we get $f^{(P)}(a', s'|a, s) = f^{(P)}(a'|a)f^{(P)}(s'|s)$ where

$$f^{(P)}(a'|a) = \frac{(\beta + \mu)\lambda e^{-\lambda a'}\left(e^{(\lambda-\beta-\mu)a\wedge a'} - 1\right)}{(\lambda - \beta - \mu)(1 - e^{-(\beta+\mu)a})}, \quad (8)$$

$$f^{(P)}(s'|s) = \frac{s' f_S(s')}{\mu_S}, \quad (9)$$

i.e. the same density with respect to age but different with respect to social index.

3.2 Phase transition

Here we shall apply the results of Bollobás, Janson and Riordan (2007) to the introduced above models in the limit as $t \to \infty$. Therefore first we shall put our model into a general setup of the theory (and notation) of inhomogeneous random graphs of Bollobás, Janson and Riordan (2007).

Let us enumerate the individuals at time t by $i = 1, \ldots, n := Y(t)$. Let $X_i = (A_i, S_i)$ denote, correspondingly, the age and social index of individual i. Notice, that this means that this individual was born at time $t - A_i$ (and remains alive at time t). The random variables X_i are independent for different i. For a randomly selected i, as we argued above, the distribution of the social index follows the original social index distribution and A_i is $\mathrm{Exp}(\gamma)$; denote the corresponding measures by μ_1 and μ_2:

$$\begin{aligned} \mu_1(da) &= \lambda e^{-\lambda a} da, \\ \mu_2(ds) &= f_S(s) ds. \end{aligned} \tag{10}$$

We shall call X_i the type of individual i. Given the types $X_i = (a_i, s_i)$ and $X_j = (a_j, s_j)$ of individuals i and j, we derive the probability $p_{ij}(n)$ that they are connected. From above we know that the expected number of (a', s')-individuals our i is connected to equals $m(a', s'|a_i, s_i) da' ds'$ defined by (4) for the U-version (and similarly for the P-version). In total, if $Y(t) = n$, the expected total number of individuals of type (a', s') equals $n \lambda e^{-\lambda a'} da' f_S(s') ds'$. This reasoning implies that the probability that i is connected to j (of type (a_j, s_j)) equals the ratio of these two expressions:

$$p_{ij}(n) = \frac{1}{n} \kappa_1(a_i, a_j) \, \kappa_2(s_i, s_j), \tag{11}$$

where

$$\kappa_1(a_i, a_j) = \frac{e^{(\lambda - \beta - \mu)(a_i \wedge a_j)} - 1}{\lambda - \beta - \mu}.$$

and

$$\kappa_2(s_i, s_j) = \begin{cases} \kappa_2^U(s_i, s_j) = \alpha(s + s') & \text{in U-version}, \\ \\ \kappa_2^P(s_i, s_j) = 2\alpha ss' & \text{in P-version}. \end{cases}$$

Let $\mu = \mu_1 \times \mu_2$, which by (10) is a probability measure on $\mathcal{S} = \mathbf{R}_+^2$, and let

$$\kappa = \kappa(x_i, x_j) = \kappa_1(a_i, a_j) \kappa_2(s_i, s_j).$$

Given the sequence $x_1, \ldots, x_n$, we let $G^{\mathcal{V}}(n, \kappa)$ be the random graph on $\{1, \ldots, n\}$, such that any two vertices i and j are connected by an edge, independently of the others vertices, with probability given by (11), which we write as

$$p_{ij}(n) = \min\{\kappa(x_i, x_j)/n, 1\}.$$

Hence, conditionally on $Y(t) = n$ the graph $G^{\mathcal{V}}(n, \kappa)$ describes our model. On the other hand, $G^{\mathcal{V}}(n, \kappa)$ with κ as above satisfies the definition of inhomogeneous random graph from Bollobás, Janson and Riordan (2007), and thus we can apply some known results here.

Recall first the fundamental result on Phase Transitions in the inhomogeneous random graphs. Define

$$T_\kappa f(x) = \int_S \kappa(x, y) f(y) d\mu(y),$$

and

$$\|T_\kappa\| = \sup\{\|T_\kappa f\|_2 : f \geq 0, \|f\|_2 \leq 1\}.$$

Then, by Theorem 3.1 from [3], the largest connected component of $G^\mathcal{V}(n, \kappa)$, denoted $C_1(G^\mathcal{V}(n, \kappa))$, satisfies

$$\frac{1}{n} C_1(G^\mathcal{V}(n, \kappa)) \xrightarrow{P} \rho_\kappa = \int_S \rho(x) d\mu(x), \text{ as } n \to \infty,$$

where $\rho_\kappa > 0$ if and only if $\|T_\kappa\| > 1$.

This gives as well for our model that conditionally on the number $Y(t)$ of individuals in the population at time t, we have with a high probability approximately $\rho_\kappa Y(t)$ individuals connected; and thus this holds unconditionally as well. The function $\rho_\kappa(x)$ is described more precisely in the following theorem.

Theorem 3.1 ([3], Theorem 6.1) *Suppose that κ is the kernel on $(\mathcal{S}, \mu)$, that $\kappa \in L^1$, and*

$$\int_\mathcal{S} \kappa(x, y) d\mu(y) < \infty$$

for every $x \in S$. Then $\rho_\kappa(x)$ is the maximal solution to

$$f(x) = 1 - \exp\{-T_\kappa f(x)\}. \tag{12}$$

Furthermore:
(i) If $\|T_\kappa\| \leq 1$ then $\rho_\kappa(x) = 0$ for every x, and (12) has only the zero solution.
(ii) If $1 < \|T_\kappa\| \leq \infty$ then $\rho_\kappa(x) > 0$ on a set of a positive measure. If, in addition, κ is irreducible, i.e., if

$$A \subseteq S \text{ and } \kappa = 0 \text{ a.e. on } A \times (S \setminus A) \text{ implies } \mu(A) \text{ or } \mu(S \setminus A) = 0,$$

then $\rho_\kappa(x) > 0$ for a.e. x, and $\rho_\kappa(x)$ is the only non-zero solution of (12).

First we shall derive conditions for $\rho_\kappa(x) > 0$. By the cited results this amounts to computing $\|T_\kappa\|$. Unfortunately, there is no general formula to simply compute the norm $\|T_\kappa\|$. Therefore we shall use a criteria for the condition when $\|T_\kappa\| > 1$, which is what we need here.

Proposition 3.2 *[[3], Proposition 17.2] For $k \geq 1$ let*

$$\alpha(k) := \frac{1}{2} \int_{\mathcal{S}^{k+1}} \kappa(x_0, x_1) \kappa(x_1, x_2) \dots \kappa(x_{k-1}, x_k) d\mu(x_0) \dots d\mu(x_k).$$

Then $\|T_\kappa\| > 1$ if and only if $\alpha(k) \to \infty$ as $k \to \infty$.

It follows from definition (10) that for our model

$$\alpha(k) = \frac{1}{2}\alpha_1(k)\alpha_2(k), \tag{13}$$

where for $i = 1,\ 2$,

$$\alpha_i(k) = \int_{\mathbf{R}_+^{k+1}} \kappa_i(u_0, u_1)\kappa_i(u_1, u_2)\ldots\kappa_i(u_{k-1}, u_k)d\mu_i(u_0)\ldots d\mu_i(u_k). \tag{14}$$

First we shall compute $\alpha_2(k)$. We start with the U-version in the next Proposition.

Proposition 3.3 *Introduce the matrix*

$$A = \left(\begin{array}{cc} \mathsf{E}[S] & 1 \\ \mathsf{E}[S^2] & \mathsf{E}[S] \end{array}\right).$$

Then for all $k \geq 3$

$$\alpha_2^U(k) = \alpha^k \left\{ A^{k-1}\left(\begin{array}{c} 2\mathsf{E}[S] \\ (\mathsf{E}[S])^2 + \mathsf{E}[S^2] \end{array}\right)\right\}_{11}, \tag{15}$$

where for a matrix B we denote $\{B\}_{ij}$ the corresponding entry.

Proof. Let $S_0, S_1, \ldots,$ be *i.i.d.* copies of the random variable S. Then by the definition (14)

$$\alpha_2^U(k) = \alpha^k \mathsf{E}\left[\prod_{i=1}^k (S_{i-1} + S_i)\right] =: \alpha^k f(k), \tag{16}$$

where

$$f(k) = \mathsf{E}\left[\prod_{i=1}^k (S_{i-1} + S_i)\right].$$

Denote also

$$g(k) = \mathsf{E}\left[\left(\prod_{i=1}^k (S_{i-1} + S_i)\right) S_k\right].$$

Then we recursively derive

$$f(k) = g(k-1) + f(k-1)\mathsf{E}[S],$$

and

$$g(k) = g(k-1)\mathsf{E}[S] + f(k-1)\mathsf{E}[S^2],$$

for $k > 1$, with

$$f(1) = 2\mathsf{E}[S], \qquad g(1) = (\mathsf{E}[S])^2 + \mathsf{E}[S^2]. \tag{17}$$

Hence, for all $k > 1$

$$\left(\begin{array}{c} f(k) \\ g(k) \end{array}\right) = A\left(\begin{array}{c} f(k-1) \\ g(k-1) \end{array}\right) = A^{k-1}\left(\begin{array}{c} f(1) \\ g(1) \end{array}\right),$$

and formula (15) follows from here, (17) and (16). $\qquad\square$

Corollary 3.4 *One has*

$$\lim_{k\to\infty} \left(\alpha_2^U(k)\right)^{1/k} = \alpha \left(\mathsf{E}[S] + \sqrt{\mathsf{E}[S^2]}\right). \tag{18}$$

In the case of the P-version, when $\alpha_2 = \alpha_2^P$ it is straightforward to derive from definition (14) that

$$\alpha_2^P(k) = (2\alpha)^k \left(\mathsf{E}[S]\right)^2 \left(\mathsf{E}[S^2]\right)^{k-2}, \tag{19}$$

which yields

$$\lim_{k\to\infty} \left(\alpha_2^P(k)\right)^{1/k} = 2\alpha\mathsf{E}[S^2]. \tag{20}$$

Define now

$$c^{cr}(\lambda, \mu + \beta) := \sup\{x > 0 : \sum_{k=2}^{\infty} x^k \, \alpha_1(k) < \infty\}. \tag{21}$$

Then due to the definitions (13) and (14) together with asymptotics (18) and (20) we have the following criteria which holds for both cases, P-version and U-version.

Corollary 3.5 *If*

$$\lim_{k\to\infty} \left(\alpha_2(k)\right)^{1/k} > c^{cr}(\lambda, \mu + \beta)$$

then $\alpha(k) \to \infty$; if

$$\lim_{k\to\infty} \left(\alpha_2(k)\right)^{1/k} < c^{cr}(\lambda, \mu + \beta)$$

then $\alpha(k) < \infty$.

The value $c^{cr}(\lambda, \mu + \beta)$ is known from [11]. Let us record this result here. Write first

$$\alpha_1(k) = \int_{\mathbf{R}_+^{k+1}} \kappa_1(a_0, a_1) \cdot \ldots \cdot \kappa_1(a_{k-1}, a_k)\lambda e^{-\lambda a_0} da_0 \ldots \lambda e^{-\lambda a_k} da_k$$

$$= \mathsf{E}\left[\kappa_1\left(\frac{X_0}{\lambda}, \frac{X_1}{\lambda}\right) \cdot \ldots \cdot \kappa_1\left(\frac{X_{k-1}}{\lambda}, \frac{X_k}{\lambda}\right)\right],$$

where $X_0, \ldots, X_k$ are *i.i.d.* random variables with a common Exp (1)-distribution. Then by the results of [12] (see [12] formula (1.7), or consult [3]) the value $c^{cr}(\lambda, \mu + \beta)$ is the smallest positive root of

$$H(x) = 1 + \sum_{n=1}^{\infty} (-1)^n \left(\frac{x}{\mu + \beta}\right)^n \frac{1}{n!} \left[\prod_{l=1}^{n} \frac{1}{1 + (l-1)\frac{\mu+\beta}{\lambda}}\right]. \tag{22}$$

Finally, combining Corollary 3.5 and Proposition 3.2 together with asymptotics (18) and (20) we can derive, apart from the critical case, the necessary and sufficient conditions for the existence of the giant component.

Corollary 3.6 *Denote*

$$\kappa^U = \kappa_1\kappa_2^U, \quad \kappa^P = \kappa_1\kappa_2^P,$$

and let

$$R^U(\alpha, \beta, \lambda, \gamma, S) := \frac{\alpha\left(\mathsf{E}[S] + \sqrt{\mathsf{E}[S^2]}\right)}{c^{cr}(\lambda, \mu + \beta)} \quad and \quad R^P(\alpha, \beta, \lambda, \gamma, S) := \frac{2\alpha\mathsf{E}[S^2]}{c^{cr}(\lambda, \mu + \beta)}. \tag{23}$$

Then there exists a giant component in the U-model if $R^U(\alpha, \beta, \lambda, \gamma, S) > 1$, and there is no giant component if $R^U(\alpha, \beta, \lambda, \gamma, S) < 1$.

Similarly, there exists a giant component in the P-model if $R^P(\alpha, \beta, \lambda, \gamma, S) > 1$, and there is no giant component if $R^P(\alpha, \beta, \lambda, \gamma, S) < 1$.

Note also that although we define the function $c^{cr}(\lambda, u)$ for $u > 0$, the model with $\mu + \beta = 0$ is treated in a very similar way, in fact when also $S \equiv 1$, this model becomes the one studied by Callaway *et al.* (2001). Then one defines $c^{cr}(\lambda, 0)$ as the critical parameter, above which there is a giant component. It is known (see Callaway *et al.* (2001)) that $c^{cr}(\lambda, 0) = 1/4$, and moreover function $c^{cr}(\lambda, u)$ is continuous at $u = 0$ (see Bollobás *et al.* (2007) and Turova (2007a)).

In general, there is no closed form for the function c^{cr}, but one can mention some qualitative properties of this function. First of all, simply from the model it follows that $c^{cr}(\lambda, \mu + \beta)$ is decreasing in λ and increasing in $\mu + \beta$, i.e., both in μ and β. For any $u > 0$ denote $c(u)$ the smallest positive root of the rescaled function (22)

$$H_u(x) = 1 + \sum_{n=1}^{\infty} (-1)^n x^n \frac{1}{n!} \left[\prod_{l=1}^{n} \frac{1}{1 + (l-1)u} \right].$$

Then

$$c^{cr}(\lambda, \mu + \beta) = (\mu + \beta) c\left(\frac{\mu + \beta}{\lambda} \right).$$

3.3   Degree correlation

We now derive an expression for the degree correlation ρ (also known as "mixing coefficient") of the network when it has grown and reached its stationary phase. That is, after a long time t we take a snap-shot of the network and compute the degree correlation ρ, which is defined as the correlation of the degrees of the two adjacent nodes of a randomly selected edge in the network. We derive ρ in two different ways. In the first method we first pick a random *node* in the snap-shot network and then perform a random walk on this fixed network. When this random walk has reached stationarity we consider the node in one step as the "first" node and then pick the "second" node randomly among the neighbouring nodes, and let the edge between these two nodes be our randomly selected edge. In the second method we pick our "first" node randomly in a size-biased way in the snap-shot network (such that the node is adjacent to a randomly selected *edge*) and the "second" node randomly among the neighbours of this individual (the size-biased distribution of a non-negative random variable X with density/pdf $f_X(x)$ and finite mean μ_X has density/pdf $x f_X(x)/\mu_X$ – below we denote such a random variable by $\tilde{X}$). As will be seen, the two methods of picking neighbouring nodes give the same result. The degree correlation is then obtained by computing the degree correlation of the two neighbouring nodes.

3.3.1   The type-distribution obtained using random walk method

We start with the first method where we first pick a node at random among all living nodes. From the results of Section 2.3 this means the node has exponential age distribution and independent social index distribution given by the original distribution. More precisely

the type distribution of a randomly selected node equals $\lambda e^{-\lambda a} da f_S(s) ds$. The neighbours of this node, conditional on its type, all have distribution $f(a', s'|a, s)$, defined in Section 3.1, where its form depends on whether we consider the U-version or the P-version of the model. Iterating this procedure gives the type-distribution of nodes after additional steps in the random walk. Eventually this random walk reaches stationarity and the distribution $f_\infty(a', s')$ of the type of node is then given by the solution to the functional equation

$$f_\infty(a', s') = \int \int f_\infty(a, s) f(a', s'|s, a) ds da.$$

And, since $f(a', s'|s, a) = f(a'|a) f(s'|s)$ in both versions of the model, it follows that the age and social index distributions are independent in the stationary distribution as well: $f_\infty(a', s') = f_\infty(a') f_\infty(s')$. The corresponding densities are hence the solutions to

$$f_\infty(a') = \int f_\infty(a) f(a'|a) da, \tag{24}$$

$$f_\infty(s') = \int f_\infty(s) f(s'|s) ds. \tag{25}$$

The functions $f(a'|a)$ and $f(s'|s)$ were defined in Section 3.1, $f(a'|a)$ being the same for the two version and $f(s'|s)$ being different for the U-version and the P-version. The (unique) solution $f_\infty(a')$ to (24) is given by

$$f_\infty(a') = \lambda \left(1 + \frac{\lambda}{\beta + \mu}\right) e^{-\lambda a'} \left(1 - e^{-(\beta+\mu)a'}\right). \tag{26}$$

As for the solution to (25), its solution depends on which version of the model we consider, i.e. whether we use $f^{(P)}(s'|s)$ or $f^{(U)}(s'|s)$. The (unique) solutions $f_\infty^{(P)}(s')$ and $f_\infty^{(U)}(s')$ are given by

$$f_\infty^{(P)}(s') = s' f_S(s')/\mu_S, \tag{27}$$

and

$$f_\infty^{(U)}(s') = \frac{s' + \mu_S}{2\mu_S} f_S(s'). \tag{28}$$

The P-version hence has the size-biased version of the social index distribution whereas the U-version has a mixture of the original and the size-biased social index distribution.

The stationary distribution defined in (26) and (27) for the P-version, and (26) and (28) for the U-version, is hence the distribution of the "first" node of a randomly selected edge. Given the type (a', s') of this individual, a randomly selected neighbour of this individual has the previously defined type-distribution $f(a, s|a', s') = f(a|a') f(s|s')$ (where $f(s|s')$, but not $f(a|a')$ depends on which model version we consider).

3.3.2 The type-distribution obtained using the size-biased method

We now derive the type distribution of the nodes of a randomly selected edge using the second method described above, which will be seen to give the same solution. We do this for the P-version of the model.

We know from before (see Equation 2) that the degree D of a randomly selected node in the snap-shot network is mixed Poisson with (random) parameter

$$2\alpha S(1 - e^{-(\beta+\mu)A})/(\beta + \mu), \tag{29}$$

where S has the original social index distribution f_S and A is independent $\text{Exp}(\lambda)$. We now instead pick our "first" node as a node of a randomly selected *edge*. As a consequence, this node will have the size-biased version $\tilde{D}$ of this random variable (because a node of degree k has probability proportional to k of being selected). The size-biased distribution of mixed Poisson is also mixed Poisson, but with the size-biased version of the random parameter. And, the size-biased distribution of the random parameter (29) equals

$$\frac{2\alpha}{\beta + \mu} \tilde{S} \left(1 - \widetilde{e^{-(\beta + \mu)A}} \right).$$

By this tilde-notation we mean that the type (A^*, S^*) of the "first" node will have social index distribution $\tilde{S}$ (the size biased distribution of the original density, having density $s f_S(s)/\mu_S$) which hence agrees with (27). The age distribution A^* will be independent and have the distribution defined by $1 - e^{-(\beta + \mu)A^*}$ being distributed like $\left(1 - \widetilde{e^{-(\beta + \mu)A}} \right)$, where $A \sim \text{Exp}(\lambda)$. But, since $A \sim \text{Exp}(\lambda)$ it follows that

$$1 - e^{-(\beta + \mu)A} \sim \text{Beta}\left(1, \ \frac{\lambda}{\beta + \mu} \right).$$

¿From this it follows that the size biased distribution

$$\left(1 - \widetilde{e^{-(\beta + \mu)A}} \right) \sim \text{Beta}\left(2, \ \frac{\lambda}{\beta + \mu} \right).$$

So $1 - e^{-(\beta + \mu)A^*}$ has this distribution, implying that $e^{-(\beta + \mu)A^*} \sim \text{Beta}\left(\lambda/(\beta + \mu), 2 \right)$. From this it follows after some straightforward calculations that A^* has exactly the density given in Equation (26). As in the previous sub-section, given the type (a', s') of the "first" node, a randomly selected neighbour of this node has the type-distribution $f(a, s|a', s') = f(a|a')f(s|s')$.

3.3.3 The degree correlation ρ

In the two previous sub-sections we have derived the type-distribution of the two nodes of a randomly selected edge: The type (A^*, S^*) of the "first" node has density $f_\infty(a')f_\infty(s')$ where $f_\infty(a')$ was defined in (26) for both versions of the model, and $f_\infty(s')$ was given by (27) for the P-version (and not explicit for the U-version). Given the type (a', s') of the first node, the second node has type-distribution $f(a, s|a', s') = f(a|a')f(s|s')$, where $f(a|a')$ and $f(s|s')$ were defined in (6) – (9). We now derive the degree correlation between two such randomly selected nodes.

In order to compute the degree correlation of the nodes adjacent to the randomly selected edge we condition on the types. Given the types we know from before that the degree distribution is Poisson: an individual of type (a, s) has a Poisson number of neighbours and the mean $g(a, s) = g_1(a)g_2(s)$ equals

$$g_1(a) = \frac{1 - e^{-(\beta + \mu)a}}{\beta + \mu} \quad \text{and} \quad g_2(s) = \begin{cases} g_2^{(U)}(s) & = \alpha(s + \mu_S) & \text{for the U-version,} \\ g_2^{(P)}(s) & = 2\alpha s & \text{for the P-version.} \end{cases}$$
$$(30)$$

We now compute the degree correlation using these results. Let D_1 denote the degree of the "first" node and D_2 to its selected neighbour. In order to compute the degree

correlation $\rho(D_1, D_2)$ we need to compute the following expectations: $\mathsf{E}[D_1] = \mathsf{E}[D_2]$, $\mathsf{E}[D_1^2] = \mathsf{E}[D_2^2]$ and $\mathsf{E}[D_1 D_2]$ (the first as well as the second moment of D_1 and D_2 are the same since we have reached stationarity). We do this excluding the edge connecting the two nodes (subtracting the value 1 everywhere does not change the degree correlation). The conditional degree distributions remain Poisson with the same means because if $X \sim \mathrm{Po}(\beta)$, then the conditional distribution of $X - 1$, conditional on that $X \geq 1$, is also $\mathrm{Po}(\beta)$. We hence get

$$\mathsf{E}[D_1] = \mathsf{E}[D_2] = \int g_1(a)f_\infty(a)da \int g_2(s)f_\infty(s)ds,$$

$$\mathsf{E}[D_1^2] = \mathsf{E}[D_2^2] = \int\int (g_1(a)g_2(s) + g_1^2(a)g_2^2(s))f_\infty(a)f_\infty(s)dads,$$

$$\mathsf{E}[D_1 D_2] = \int\int g_1(a)g_1(a')f_\infty(a)f(a'|a)dada' \int\int g_2(s)g_2(s')f_\infty(s)f(s'|s)dsds'.$$

Given these expressions, the degree correlation is given by

$$\rho(D_1, D_2) = \frac{\mathsf{E}(D_1 D_2) - \mathsf{E}(D_1)\mathsf{E}(D_2)}{\mathsf{E}(D_1^2) - (\mathsf{E}(D_1))^2} = \frac{\mathsf{C}(D_1, D_2)}{\mathsf{V}[D_1]}. \tag{31}$$

The perhaps most important questions is to learn if ρ, or equivalently the covariance C, is positive or negative. For this it is sufficient to compute the numerator of (31). Using $f_\infty(a, s)$ (defined by (26), (27), and (28)) and (30) in the expressions above for the P-version and U-version, separately, standard but tedious calculations reveal that

$$\mathsf{E}[D_1^{(P)}] = \mathsf{E}[D_2^{(P)}] = \frac{2}{\lambda + 2(\beta + \mu)}\frac{2\alpha\mathsf{E}[S^2]}{\mathsf{E}[S]},$$

$$\mathsf{E}[D_1^{(P)} D_2^{(P)}] = \frac{5\lambda + 6(\beta + \mu)}{(\lambda + \beta + \mu)(\lambda + 2(\beta + \mu))(\lambda + 3(\beta + \mu))}\left(\frac{2\alpha\mathsf{E}[S^2]}{\mathsf{E}[S]}\right)^2,$$

as well as

$$\mathsf{E}[D_1^{(U)}] = \mathsf{E}[D_2^{(U)}] = \frac{2}{\lambda + 2(\beta + \mu)}\frac{\alpha\left(\mathsf{E}[S^2] + 3(\mathsf{E}[S])^2\right)}{2\mathsf{E}[S]},$$

$$\mathsf{E}[D_1^{(U)} D_2^{(U)}] = \frac{5\lambda + 6(\beta + \mu)}{(\lambda + \beta + \mu)(\lambda + 2(\beta + \mu))(\lambda + 3(\beta + \mu))}2\alpha^2\left(\mathsf{E}[S^2] + (\mathsf{E}[S])^2\right).$$

This gives us for the P-version

$$\mathrm{C}^{(P)}(D_1, D_2) = \frac{\lambda^2}{(\lambda + \beta + \mu)(\lambda + 2(\beta + \mu))^2(\lambda + 3(\beta + \mu))}\left(2\alpha\mathsf{E}(S^2)\right)^2. \tag{32}$$

Since all parameters are positive we conclude that for the P-version of the model the covariance, and hence also the degree correlation ρ, is always positive. This is true irrespective of the model parameters and choices of social index distribution.

For the U-version the picture is different. Introduce $\gamma = \beta + \mu$ and compute

$$C^{(U)}(D_1, D_2) = \alpha^2 \frac{6\gamma + 5\lambda}{(\lambda + \gamma)(\lambda + 2\gamma)(\lambda + 3\gamma)} 2 \left(\mathsf{E}[S^2] + (\mathsf{E}[S])^2\right) \tag{33}$$

$$- \alpha^2 \frac{1}{(\lambda + 2\gamma)^2} \left(\frac{\mathsf{E}[S^2] + 3(\mathsf{E}[S])^2}{\mathsf{E}[S]}\right)^2$$

$$= \frac{\alpha^2}{(\lambda + 2\gamma)^2} (\mathsf{E}S)^2 \left(\frac{(6\gamma + 5\lambda)(\lambda + 2\gamma)}{(\lambda + \gamma)(\lambda + 3\gamma)} 2 \left(\frac{\mathsf{E}[S^2]}{(\mathsf{E}[S])^2} + 1\right) - \left(\frac{\mathsf{E}[S^2]}{(\mathsf{E}[S])^2} + 3\right)^2\right). \tag{34}$$

Let us also denote

$$a = \frac{\lambda^2}{(\lambda + \gamma)(\lambda + 3\gamma)}.$$

Then we can rewrite (33) as follows

$$C^{(U)}(D_1, D_2) = \frac{\alpha^2}{(\lambda + 2\gamma)^2} (\mathsf{E}S)^2 \left(2(a + 4)\left(\frac{\mathsf{E}[S^2]}{(\mathsf{E}[S])^2} + 1\right) - \left(\frac{\mathsf{E}[S^2]}{(\mathsf{E}[S])^2} + 3\right)^2\right) \tag{35}$$

$$= \frac{\alpha^2}{(\lambda + 2\gamma)^2} (\mathsf{E}S)^2 \left(-\left(\frac{\mathsf{E}[S^2]}{(\mathsf{E}[S])^2}\right)^2 + 2\frac{\mathsf{E}[S^2]}{(\mathsf{E}[S])^2}(a + 1) - (1 - 2a)\right).$$

This together with the facts that $\frac{\mathsf{E}[S^2]}{(\mathsf{E}[S])^2} \geq 1$ and $0 < a < 1$, yields that if

$$\frac{\mathsf{E}[S^2]}{(\mathsf{E}[S])^2} < 1 + a + \sqrt{a^2 + 4a}$$

then $C^{(U)}(D_1, D_2) > 0$, and hence also that $\rho^{(U)} > 0$. If

$$\frac{\mathsf{E}[S^2]}{(\mathsf{E}[S])^2} = 1 + a + \sqrt{a^2 + 4a}$$

then $\rho^{(U)} = 0$, and otherwise, $\rho^{(U)} < 0$.

We conclude, that for a fixed value of $\mathsf{E}[S]$ and the other model parameters, increasing the variance of S from 0 (constant S) to very large variance, allows us to pass through the assortative regime, to neutral (no assortativity) and then to disassortative. One can also observe that increasing the variance (i.e. $\mathsf{E}[S^2]$) also increase R^U and R^P respectively (see (23)) which thus might cross the critical value of 1 leading to percolation. This second effect, that disassortativity facilitates percolation, was also reported by Newman (2002) for a different model.

4 Discussion

In the present paper we studied properties of a stochastic dynamic network in which nodes are born and die randomly in time, and while alive pair of nodes are connected and disconnected by edges randomly in time. The rate of connecting to other nodes depended on the social index, given at birth, of the node. For this model we derived

limiting properties of a snapshot of the network after a long time, assuming the node population grew large.

The three main results were: the type distribution of neighbour nodes (Section 3.1), a criterion for when phase transition occurs, above which the network has a giant component (Section 3.2) and an expression for the degree correlation of connected edges (Section 3.3). One of the most interesting features of the model is that U-version of the model (in which neighbouring nodes are selected uniformly among all living nodes) may have either positive or negative degree correlation depending on the numerical values of the model parameters (birth and death rates of nodes and edge) and the social index distribution.

Together with previous results of the model (Britton and Lindholm, (2010)) the most important limiting local properties of the network are hence known by now, as well as whether or not the network has a giant component. Other global properties such as the diameter of the network remains to be analysed. Another interesting class of problems to study for the present model would be to study dynamic properties of the network. One such problem would be to study limiting properties of some process taking place "on" the network. For example, can an epidemic, forest fire or similar persist on the network forever or will it die out? What happens if a lightning process kills randomly selected nodes and all their neighbours at a constant rate.

The model can of course also be generalized in several ways to make it more applicable for certain situations. Nodes could be of different types with different attachment rates depending on the types of the nodes in question, for example male or female if mimicking a sexual network. Similarly, edges could be of different types, perhaps reflecting the "degree" of acquaintance which for example could affect the transmission probability of an epidemic taking place on the network. Still, the current model does capture some important properties of a real world network in the nodes as well as edges are created and cease to exist. Also, depending on the model parameters, the model is quite flexible in producing different degree distributions and degree correlations (the network has no clustering asymptotically).

Acknowledgments

References

[1] BARABÁSI, A.-L., ALBERT, R. (1999) Emergence of scaling in random networks. *Science*, **286**, 509–512.

[2] BARABÁSI, A.-L. ALBERT R., AND JEONG H. (1999) Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications*, **272**, 173–187.

[3] BOLLOBÁS B., JANSON S. AND RIORDAN O. (2007) The phase transition in inhomogeneous random graphs. *Random Structures and Algorithms* **31** , 3–122.

[4] BOLLOBÁS B., JANSON S. AND RIORDAN O. (2011) Sparse random graphs with clustering, *Random Structures and Algorithms*, to appear.

[5] BRITTON, T., LINDHOLM, M. (2010) Dynamic random networks in dynamic populations. *J. Stat. Phys.*, 139: 518-535.

[6] CALLAWAY, D.S., HOPCROFT, J.E., KLEINBERG, J.M., NEWMAN, M. E. J., AND STROGATZ, S. H. (2001) Are randomly grown graphs really random? *Phys. Review E*, 64 041902.

[7] HACCOU, P., JAGERS, P. AND VATUTIN, V.A. (2005) *Branching Processes: Variation, Growth, and Extinction of Populations.* Cambridge University Press, Cambridge.

[8] MALYSHEV, V. A. (1998) Random graphs and grammars on graphs. *Discrete Math. Appl.* 8:247-262.

[9] NEWMAN, M. E. J. (2002) Assortative Mixing in Networks. *Phys. Rev. Lett.* 89, 208701.

[10] TUROVA, T.S. (2002) Dynamical random graphs with memory. *Physical Review E*, **65**, 066102.

[11] TUROVA, T.S. (2003) Long Paths and Cycles in the Dynamical Graphs. *Journal of Statistical Physics*, **110**, 1/2, 385-417.

[12] TUROVA, T.S. (2007) Phase Transitions in Dynamical Random Graphs. *Journal of Statistical Physics*, **123**(5), 1007–1032.

[13] TUROVA, T.S. (2007a) Continuity of the percolation threshold in randomly grown graphs. *Electronic Journal of Probability*, 12, 1036-1047.